

# Gravity and trade in video on demand services



Annette Brooks



Zuzanna Studnicka



European Commission, Joint Research Centre (JRC)



University College Dublin

FIW-wiiw Seminars in International Economics  
28 March 2023

# Disclaimer

All errors are our own. The views expressed are purely those of the authors and may not under any circumstances be regarded as stating an official position of the European Commission.

## My List

0:00 See how we built Netflix, Pt. 1

## Introduction - Definitions

Video on demand services account for more than 60% of all internet traffic. Given the differences in the largest players' business models, it is important to define the three main types:

- ▶ subscription video on demand (SVOD): Netflix, Hulu, Disney+, HBO
- ▶ transactional video on demand (TVOD): Apple iTunes, Google Play, Sky Box Office
- ▶ ad-based video on demand (AVOD): YouTube, Pluto TV, Xumo

0:00 See how we built Netflix, Pt. 2

## Data - Countries covered

Figure: Country coverage at a glance



0:00 See how we built Netflix, Pt. 3

## Estimation strategy

We start with the following baseline gravity specification:

$$X_{ij,t}^p = \exp[\pi_{i,t} + \lambda_{j,t} + \beta_1 GDP_{i,t} + \beta_2 GDP_{j,t} + \beta_3 GDP_{i,t} + \beta_4 GDP_{j,t} + \beta_5 DIST_{ij} + \beta_6 LANG_{ij} + \beta_7 CONTIG_{ij} + \beta_8 BROAD_{i,t} + \beta_9 AWARDS_{i,t} + \sum_{g=1}^G \gamma_g genre_{ig} + \beta_{10} age_{i,t} + \beta_{11} ORIG_{i,t}] \times \epsilon_{ij,t}^p$$

where  $\pi_{i,t}$  denotes the vector of importer-time fixed effects, and  $\lambda_{j,t}$  denotes the set of exporter-time fixed effects. We define the primary production country of a title  $p$  as exporter  $j$ .

0:00 See how we built Netflix, Pt. 4

## Gravity results: baseline

Table: Baseline results: including domestic

	(1)	(2)	(3)	(4)	(5)
	baseline	age	age	age	age
logGDPi:dest	0.968	-0.508	-1.927	-1.918	-0.724
logGDP:dest	3.013	2.268	3.184	3.063	3.038
logGDPi:origin	30.566**	36.146**	38.481**	38.659**	38.893
logGDP:origin	39.596**	39.216**	39.224**	39.264**	39.708
Common language	0.866**	0.822**	0.824**	0.822**	0.822**
Contig	0.008**	0.017**	0.017**	0.016**	0.012**
Contigadj	-0.61	-0.465	-0.328	-0.398	-0.32
logBroadband	4.326**	4.244**	4.303	4.222	4.298
Woe		0.016**	0.016**	0.017**	0.017**
Children & Family		1.411**	1.411**	1.411**	1.411**
Comedy & Thriller		1.220**	1.220**	1.220**	1.220**
Crime & Fantasy		0.715	0.715	0.715	0.715
Drama		-0.026**	-0.026**	-0.026**	-0.026**
Horror		-0.170**	-0.170**	-0.170**	-0.170**
Romance		1.031**	1.031**	1.031**	1.031**
Sci-Fi & Fantasy		1.668**	1.668**	1.668**	1.668**
logAge		-1.020**	-0.863**	-0.852	-0.852
Netflix Original				2.180**	2.180**
Constant	548.048**	633.474**	577.808**	548.326**	517.463
Observations	363,586	363,586	363,586	363,586	363,586

0:00 See how we built Netflix, Pt. 5

## Robustness check, excluding US movies

Table: TV Shows vs. Movies, no US

	TVShow	All	Movie	TVShow	Foreign
	baseline	baseline	baseline	baseline	baseline
logGDPi:dest	-28.251*	-1.918	-7.298	-28.571*	16.444
logGDP:dest	79.667	2.888	2.842	78.099*	11.576
logGDPi:origin	-28.384	-89.023	-88.649	-21.849	-81.841
logGDP:origin	13.039	64.733	61.051	16.995	86.020
logDist	-0.616**	0.500**	0.500**	-0.365**	0.365**
Common language	1.399**	1.229**	0.274	1.189**	-0.892
Contigadj	0.182	0.11	0.104	-0.116	0.862
logBroadband	0.411	0.548	0.339	0.437	0.895
Woe	0.048**	0.048**	0.048**	0.048**	0.048**
Children & Family	-2.846**	2.823**	1.292*	-2.109**	1.709**
Comedy & Thriller	-0.174*	2.216**	1.614**	-0.255**	0.485
Crime & Fantasy	1.175**	1.688**	1.875	1.335**	0.339
Drama	-2.826**	2.086	0.892	-2.895**	0.369
Horror	0.616**	1.275**	0.841	0.738**	1.804
Romance	3.261**	3.023**	2.967**	3.168**	3.752
Sci-Fi & Fantasy	0.023**	0.626**	0.422**	-0.842**	4.591**
logAge	-0.828**	-1.014**	-1.269*	-0.797**	-2.229**
Netflix Original	3.070**	3.672**	3.256**	3.616**	2.677**
Constant	467.858	1,347.489	1,356.702	612.294	0,362*
Observations	468,268	468,268	468,268	468,268	1,088,143

0:00 See how we built Netflix, Pt. 6

## Conclusion

- ▶ Gravity holds on the intensive margin as well as on the extensive margin
- ▶ Being a co-produced movie in which the primary production country is not the destination country does not significantly increase clicks

# Outline of the talk

Introduction

Data

Estimation Strategy

Results

Robustness Checks

Conclusion

# Preview

## Motivation

- ▶ Streaming has increased in importance over the past 15 years
- ▶ Catalogues are fragmented and access to foreign catalogues is blocked
- ▶ Policy and legal discussion of Copyright versus Digital Single Market

## Research questions

- ▶ How fragmented are Netflix catalogues?
- ▶ What are the patterns of content availability (extensive margin) and viewing (intensive-margin)?
- ▶ Can they be explained by similar factors that drive international trade (gravity equation)?

# Preview (cont'd)

## Data

- ▶ Bi-annual number of clicks per title
- ▶ Complete catalogue data

## Overview of Results

- ▶ Netflix catalogues are very fragmented, the mean overlap of movie catalogues amounts to 63%
- ▶ Gravity controls are more important for the intensive than for the extensive margin
- ▶ Trade costs still matter even on Netflix
- ▶ Strong preference for domestic content

## Introduction - Motivation

- ▶ Intermediary connecting content producers with content consumers, including viewers in countries other than where content is produced.
- ▶ From the international trade perspective: reduces transport costs.
- ▶ But we cannot talk about frictionless trade: national subsidies in the audiovisual sector, complex dubbing license requirements etc.
- ▶ Geo-blocking still present in the EU, excluded from the Digital Single Market Strategy.
- ▶ This leads to the fragmentation of Netflix catalogues across countries.
- ▶ Netflix has more productions than big European broadcasters (and almost a monopoly to impose its content - 30% supposed to be European).

## Introduction - Definitions

Video on demand services account for more than 60% of all internet traffic. Given the differences in the largest players' business models, it is important to define the three main types:

- ▶ subscription video on demand (SVOD): Netflix, Hulu, Disney+, HBO max
- ▶ transactional video on demand (TVOD): Apple iTunes, Google Play, Sky Box Office
- ▶ ad-based video on demand (AVOD): YouTube, Pluto TV, Xumo

By now, many services have mixed models, such as Netflix, Disney+, Amazon Prime, and Peacock



## Introduction - Netflix

- ▶ The largest platform with respect to coverage and subscribers
- ▶ Present in all countries except for China, North Korea, Syria, Crimea, and now Russia
- ▶ Accounts for around 13% of global internet downstream traffic
- ▶ 230M subscribers globally, 76M in Europe, Middle East and Africa (Q4 2022)
- ▶ 58M in Europe (The Economist, April 2021)
- ▶ Increasing amount of Netflix Original content, 2.4\$bn spent in 2019, 750£m in 2020 only in UK-produced shows
- ▶ Producer/distributor (push for their own products? US-origin products?)
- ▶ Mostly attracts a younger audience than linear TV

# Introduction - Research questions

Research questions:

- ▶ How fragmented is the Netflix catalogue?
- ▶ What are the patterns of content availability (extensive margin) and content viewing (intensive-margin)?
- ▶ Can they be explained by similar factors that drive international trade (gravity equation)?
- ▶ How are they affected by Netflix-produced content?
- ▶ What is the role of show/movie characteristics (quality - not reflected in price)?

# Literature

Limited number of papers in economics:

- ▶ Trade in cultural goods and services:
  - ▶ Marvasti and Canterbury (Econ. Inq., 2005), Disdier et al. (ROWE, 2010), Ferreira and Waldfogel (EJ, 2013).
- ▶ Netflix-specific:
  - ▶ Batikas et al. (JRC, 2015) analyse the availability of Netflix content in 11 EU countries.
  - ▶ Aguiar and Waldfogel (J. Cult. Econ., 2018) analyse US-origin content availability
- ▶ Importance of quality in trade:
  - ▶ Crozet et al. (RES, 2012)
  - ▶ Chen and Juvenal (JIE, 2016)

# Hypotheses

Based on our research questions and the related literature, we focus on four main hypotheses.

1. There is a large variation in the content of Netflix catalogues across countries, reflecting different distribution patterns such as copyrights requirements etc.
2. Gravity variables matter more at the intensive margin level since it reflects the audience preferences within a country.
3. Netflix original content attracts a large proportion of clicks.
4. Quality is a major determinant of Netflix viewing.

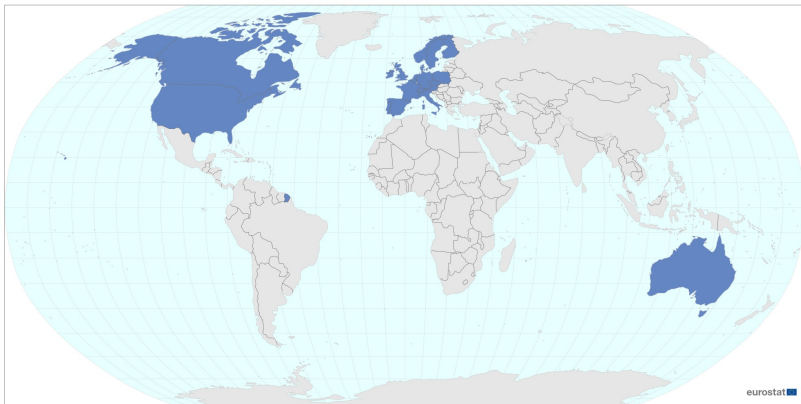
# Data

# Data

- ▶ Netflix viewing data from January 2018 to December 2019 (half-year periods)
  - ▶ Information of desktop clicks to Netflix sub-pages (titles) from SimilarWeb
- ▶ Covers 20 destination countries: 15 EU, 5 non-EU and 80 countries of origin.
- ▶ National Netflix catalogue information from Ampere Analysis.
- ▶ Scraped data to obtain info on title and year from urls such as `netflix.com/title/70254521` to a specific film or TV show title.
- ▶ Information on international co-productions from Lumiere VoD.
- ▶ IMDb ratings from IMDb.
- ▶ Other data: gravity from CEPII.

# Data - Countries covered

Figure: Country coverage at a glance



# Data - Netflix catalogue: descriptive statistics

Table: Numbers and shares

	Number	Share	Clicks	Share of clicks
All	14,379		764.3M	
2018	11,730		363M	47%
2019	11,688		401M	53%
Movies	11,656	81.06%	4.3M	0.56%
TV Shows	2,723	18.94%	760M	99.44%
Netflix Original	496	3.44%	499.8M	65.40%
Netflix Exclusive	538	3.73%	65.9M	8.60%



## Data - Netflix catalogue: descriptive statistics cont.

Table: Views by content

Genre	Number	Share	Clicks	Share clicks
Action & Adventure	1,285	8.93%	35.2M	4.6%
Children & Family	2,319	16.10%	19.4M	2.5%
Comedy	2,581	17.93%	340M	44.5%
Crime & Thriller	2,760	19.18%	122M	16.0%
Drama	1,406	9.77%	30M	3.9%
Horror	609	4.23%	21M	2.7%
Romance	2,167	15.05%	21M	2.7%
Sci-Fi & Fantasy	1,252	8.69%	174M	22.8%
Total	14,379	100	764.3M	100%

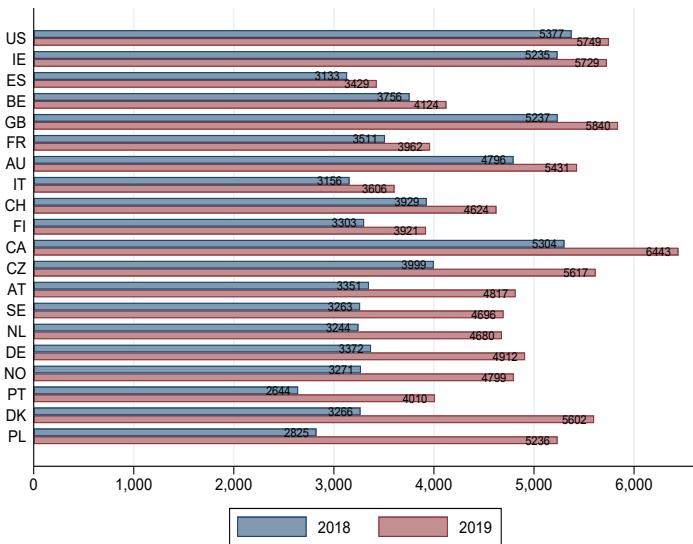
# Data - Netflix catalogue: descriptive statistics cont.

Table: Most watched content

	Title	Type	Original	Year	IMDb	Origin	Clicks (M)
1	the office (u.s.)	TVShow	No	2005	8.8	US	103.9
2	brooklyn nine-nine	TVShow	No	2013	8.4	US	32.5
3	stranger things	TVShow	Yes	2016	8.8	US	27.9
4	big mouth	TVShow	Yes	2017	8	US	20.0
5	13 reasons why	TVShow	Yes	2017	7.8	US	19.1
6	money heist	TVShow	Yes	2017	8.5	ES	18.8
7	disenchantment	TVShow	Yes	2018	7.2	US	18.3
8	love, death and robots	TVShow	Yes	2019	8.6	US	15.4
9	lucifer	TVShow	Yes	2015	8.2	US	14.6
10	you	TVShow	Yes	2018	7.8	US	14.5
90	black mirror: bandersnatch	<b>Movie</b>	Yes	2018	7.2	US	1.7
122	el camino: a breaking bad movie	<b>Movie</b>	Yes	2019	7.4	US	0.89

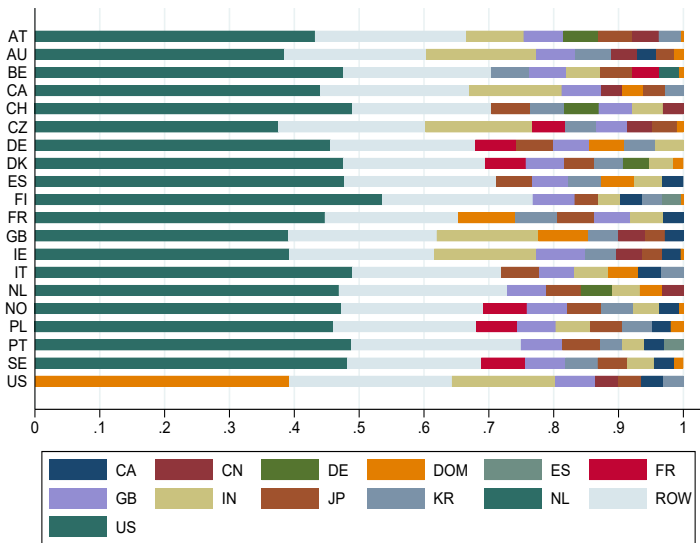
# Data - Netflix catalogue: by country

Figure: Number of titles by country



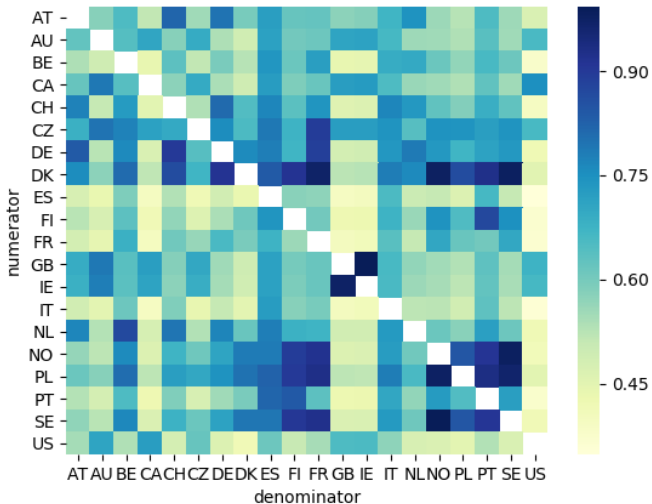
# Data - Netflix catalogue: by country

Figure: Share of different production countries, 2019



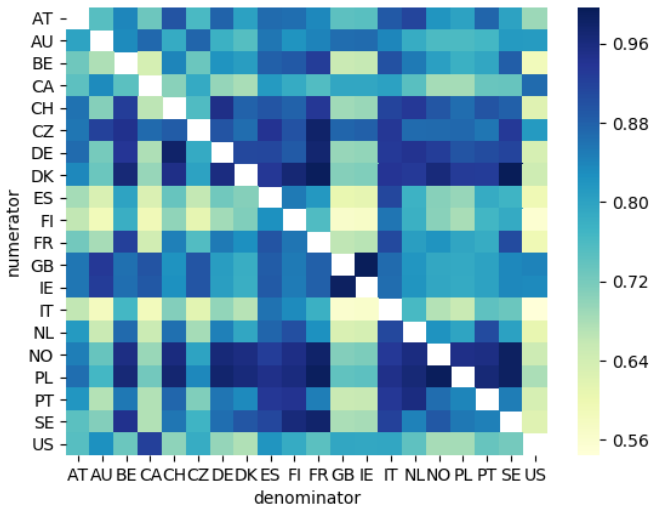
# Data - Netflix catalogue: by country

Figure: Movies catalogues overlap by country, 2019 H2



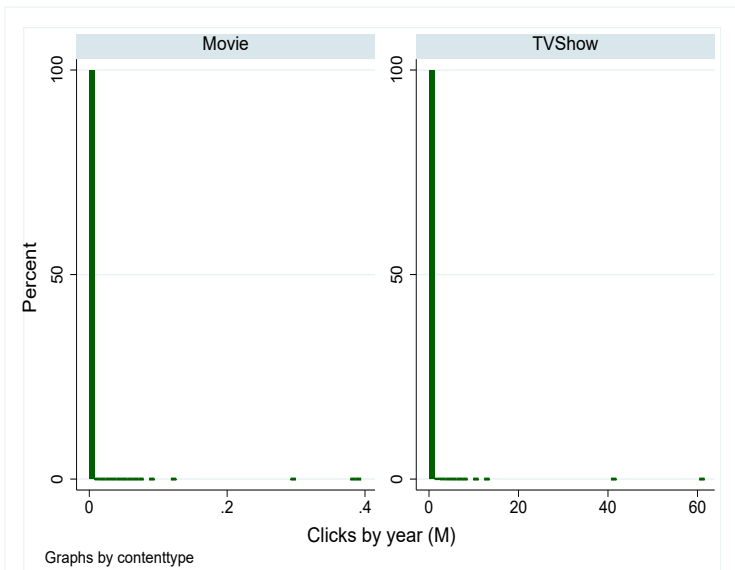
# Data - Netflix catalogue: by country

Figure: TV Shows catalogues overlap by country, 2019 H2



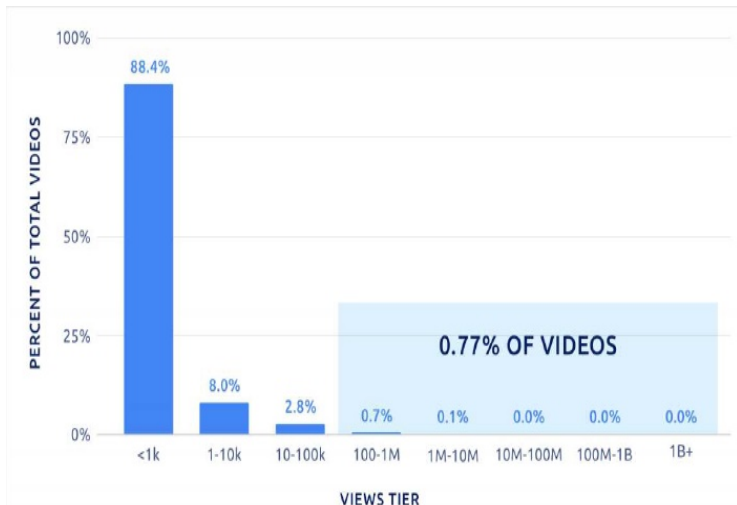
# Data - Netflix catalogue: histogram

Figure: Histogram all clicks



# Data - Comparison: YouTube clicks

Figure: Total videos by tier



Source: <https://pex.com/blog/state-of-youtube-2019-music-more-valuable>



# Summary Statistics

Table: Summary statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
Extensive margin	5,940	48.41	204.75	0.00	2,464
Clicks	203,688	3,680	137,969	0.00	33,400,000
logDist.	203,688	8.37	1.01	4.23	9.88
Domestic	203,688	0.06	0.23	0.00	1.00
Common language	203,688	0.20	0.40	0.00	1.00
Contiguity	203,688	0.07	0.26	0.00	1.00
Colony	203,688	0.13	0.33	0.00	1.00
Rating	203,688	6.53	1.17	1.40	9.50
logRating	203,688	1.86	0.20	0.34	2.25
Age	203,688	9.30	9.43	1.00	100.00
logAge	203,688	1.85	0.86	0.00	4.61
Average Length	203,688	79.91	36.47	1.00	251.00
logAv. Length	203,688	4.22	0.65	0.00	5.53
Netflix Original	203,688	0.23	0.42	0.00	1.00

# Estimation Strategy

## Estimation strategy: content availability

We begin with content availability (number of titles):

$$Y_{i,j,t} = \alpha + fe_{it} + fe_{jt} + \beta_1 x_{i,j} + \varepsilon_{i,j,t}$$

$Y_{i,j,t}$ : number of titles from country  $j$  available in country  $i$  at time  $t$

$x_{i,j}$ : gravity variables: distance, domestic dummy, stock of migrants, common language, contiguity, colony

## Estimation strategy: number of clicks

Our main specification:

$$y_{i,j,t}^p = \alpha + fe_{it} + fe_{jt} + \beta_1 x_{i,j} + \beta_2 x_p + \beta_3 age_{p,t} + \sum_{g=1}^8 \gamma_g genre_{p,g} + \varepsilon_{ij,t}^p,$$

$y_{i,j,t}$ : number clicks going to an individual title  $p$  (from country  $j$  in country  $i$ ) at time  $t$ ,

$x_{i,j}$ : gravity variables

$x_p$ : show/movie characteristics (IMDb rating, length and Netflix Original dummy)

$age_{p,t}$ : the show/movie's age

$genre_{p,g}$ : genre dummies

# Results

# Results: content availability (extensive margin)

Table: Content availability (number of titles)

	All countries			The same as in IM		Excl. US	
	(1) all	(2) all	(3) foreign	(4) all	(5) foreign	(6) all	(7) foreign
logDist.	-0.097** (0.038)	0.009 (0.042)	0.024 (0.036)	0.023 (0.037)	0.033 (0.033)	-0.116*** (0.031)	-0.069** (0.030)
Domestic		0.456*** (0.166)		0.529*** (0.155)		0.713*** (0.121)	
logMigrants			0.097*** (0.021)		0.102*** (0.023)		0.062*** (0.014)
Colony	0.046 (0.034)	0.027 (0.040)	0.013 (0.041)	0.019 (0.039)	0.010 (0.041)	-0.129** (0.051)	-0.226*** (0.060)
Common language	0.273** (0.115)	0.318*** (0.118)	0.132 (0.102)	0.354*** (0.122)	0.154 (0.104)	0.694*** (0.098)	0.564*** (0.101)
Contiguity	-0.156 (0.097)	0.043 (0.102)	0.050 (0.102)	0.071 (0.098)	0.064 (0.101)	-0.039 (0.056)	0.011 (0.061)
Constant	6.651*** (0.308)	5.726*** (0.356)	4.663*** (0.356)	5.776*** (0.311)	4.699*** (0.353)	5.347*** (0.255)	4.412*** (0.280)
Observations	5,940	5,940	5,628	2,480	2,412	5,860	5,552
Pseudo R2	0.956	0.956	0.960	0.949	0.955	0.903	0.907

PPML estimation results. Dependant variable: number of titles available. All results include origin-time, destination-time and time fixed effects. Standard errors are clustered at the origin-time and destination-time level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

# Results: number of clicks (intensive margin)

Table: Baseline results, TV Shows

	(1)	(2)	(3)	(4)	(5)	(6)
logDist.	-0.469***	-0.262	-0.324*	-0.336*	-0.305*	-0.306*
Domestic		0.666	0.582	0.647	0.753	0.756*
Language	1.402***	1.498***	1.585***	1.524***	1.519***	1.513***
Contiguity	-0.604	-0.181	-0.305	-0.278	-0.158	-0.163
Colony	-0.358	-0.372	-0.364	-0.317	-0.296	-0.288
logRating			10.197***	11.931***	11.355***	11.305***
logAge				-0.791***	-0.501***	-0.493***
Children & Family				-1.740***	-1.728***	-1.997***
Comedy				0.640***	0.616***	0.438**
Crime & Thriller				-0.597***	-0.561***	-0.546***
Drama				-0.896***	-0.834***	-0.767***
Horror				-0.210**	-0.192*	-0.253**
Romance				-0.335	-0.171	-0.241
Sci-Fi & Fantasy				0.356***	0.383***	0.314***
Netflix Original					0.790***	0.821***
logAv.Length						-0.344***
Constant	14.151***	12.198***	-8.098***	-10.356***	-10.393***	-8.999***
Observations	68,432	68,432	68,432	68,432	68,432	68,432
Pseudo R2	0.319	0.319	0.423	0.510	0.519	0.520

PPML estimation results. Dependant variable: number of clicks. All results include origin-time destination-time fixed effects. Standard errors are clustered at title level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

# Results: number of clicks (intensive margin)

Table: Foreign TV Shows

	(1)	(2)	(3)	(4)	(5)	(6)
logDist.	-0.198**	-0.272***	-0.292***	-0.296***	-0.285***	-0.283***
logMigrants		0.417***	0.409***	0.426***	0.445***	0.440***
Common language	1.514***	1.467***	1.534***	1.526***	1.562***	1.552***
Contiguity	0.055	-0.656**	-0.672**	-0.633**	-0.590**	-0.582**
Colony	-0.473***	-0.396***	-0.425***	-0.363***	-0.327***	-0.324***
logRating			8.610***	10.129***	9.706***	9.704***
logAge				-0.976***	-0.648***	-0.646***
Children & Family				-2.012***	-1.998***	-2.231***
Comedy				0.258	0.243	0.089
Crime & Thriller				-0.570*	-0.552*	-0.538
Drama				-0.902	-0.862	-0.745
Horror				-0.364	-0.382	-0.433
Romance				-0.151	0.011	-0.052
Sci-Fi & Fantasy				0.307	0.351	0.287
Netflix Original					0.937***	0.956***
logAv.Length						-0.310
Constant	10.848***	6.785***	-10.485***	-12.173***	-12.747***	-11.533***
Observations	64,809	64,809	64,809	64,809	64,809	64,809
Pseudo R2	0.246	0.248	0.344	0.452	0.465	0.466

PPML estimation results. Dependant variable: number of clicks. All results include origin-time, destination-time fixed effects. Standard errors are clustered at the origin-time, destination-time. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .



# Results: number of clicks (intensive margin)

Table: Movies

	(1)	All (2)	(3)	(4)	Foreign (5)	(6)
logDist.	0.224	0.227	0.139	0.215	0.216	0.173
Domestic	1.801*	1.805*	1.206			
logMigrants				0.483***	0.483***	0.529***
Common language	1.085	1.086	0.719	0.366	0.369	0.034
Contiguity	1.281	1.290	1.167	0.048	0.049	0.020
Colony	0.154	0.158	0.164	-0.322	-0.315	-0.330
logRating	9.030***	9.359***	9.285***	10.162***	10.456***	10.213***
logAge	-1.790***	-1.791***	-1.783***	-1.852***	-1.846***	-1.780***
Children & Family	1.251	1.028	0.985	0.354	0.191	0.224
Comedy	-0.113	-0.248	-0.269	-0.943***	-1.025***	-0.936**
Crime & Thriller	1.991*	1.956*	1.904*	0.808	0.826	0.897
Drama	-1.278	-1.322	-1.360	-2.414**	-2.398**	-2.270**
Horror	-0.534	-0.541	-0.603	-1.777***	-1.739***	-1.706***
Romance	1.075	1.007	0.958	0.152	0.134	0.208
Sci-Fi & Fantasy	3.134**	3.076***	2.944***	2.226***	2.217***	2.177***
Netflix Original	1.655***	1.654***	1.632***	1.544***	1.545***	1.540***
logLength		-0.329	-0.294		-0.317	-0.273
Co-production			0.858*			1.067***
Constant	-14.650***	-13.720***	-12.873***	-20.722***	-19.822***	-19.852***
Observations	135,256	135,256	135,256	123,207	123,207	123,207
Pseudo R2	0.581	0.582	0.586	0.602	0.602	0.607

PPML estimation results. Dependant variable: number of clicks. All results include origin-time and destination-time fixed effects. Standard errors are clustered at the origin-time destination-time level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

# Results: number of clicks (intensive margin)

Table: By Netflix Original, TV Shows

	All		Foreign	
	(1) not nfo	(2) nfo	(3) not nfo	(4) nfo
logDist.	-0.805***	-0.187**	-1.093***	-0.246***
Domestic	3.342***	1.008***		
logMigrants			0.472**	0.530***
Common language	2.912***	1.256***	3.564***	1.326***
Contiguity	-1.218	0.322	-1.572	-0.511*
Colony	0.407	-0.153	0.088	-0.120
logRating	26.007***	6.878***	14.642***	7.048***
logAge	0.284**	-0.869***	-0.162	-0.914***
Children & Family	-0.458	-2.491***	-0.804	-2.775***
Comedy	2.737***	-0.407	2.241***	-0.466
Crime & Thriller	-1.232**	-0.769*	-0.402	-0.780**
Drama	-1.735**	-1.015*	-2.378***	-1.086*
Horror	-1.274***	-0.490**	-0.556	-0.621**
Romance	1.293**	-1.995***	2.353***	-2.845***
Sci-Fi & Fantasy	-1.281***	0.213	-0.574*	0.246
logAv.Duration	-1.128***	0.076	-1.237***	0.102
Constant	-38.556***	-0.801	-15.662***	-7.269***
Observations	29,056	30,059	26,561	28,776
Pseudo R2	0.705	0.546	0.618	0.482
Log likelihood	-3.820e+08	-8.750e+08	-1.810e+08	-5.350e+08

PPML estimation results. Dependant variable: number of clicks. All results include origin-time and destination-time fixed effects. Standard errors are clustered at the origin-time destination-time level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

# Robustness Checks

# Robustness checks: different language measures

Table: Robustness check: Melitz and Toubal linguistic proximity measures

	(1) col	(2) csl	(3) cnl	(4) three	(5) three + lp	(6) index
logDist.	-0.284***	-0.349***	-0.297***	-0.323***	-0.349***	-0.366***
logMigrants	0.441***	0.464***	0.458***	0.449***	0.467***	0.470***
Common language	1.553***			0.731**	0.873***	
Common spoken		3.160***		1.227***	-0.375	
Common native			2.410***	0.682	2.267***	
Adjusted Ling. Prox.					0.429***	
Common lang. Index						3.037***
Contiguity	-0.584**	-0.697**	-0.623**	-0.734***	-0.878***	-0.858***
Colony	-0.325***	-0.070	-0.460***	-0.296***	-0.366***	-0.286**
logRating	9.652***	9.638***	9.652***	9.651***	9.657***	9.652***
logAge	-0.639***	-0.642***	-0.639***	-0.640***	-0.639***	-0.640***
Children & Family	-2.202***	-2.207***	-2.203***	-2.204***	-2.202***	-2.203***
Comedy	0.109	0.107	0.109	0.109	0.111	0.111
Crime & Thriller	-0.538*	-0.533	-0.538*	-0.537	-0.541*	-0.540*
Drama	-0.807	-0.806	-0.808	-0.807	-0.809	-0.809
Horror	-0.423	-0.423	-0.423	-0.423	-0.423	-0.422
Romance	-0.047	-0.037	-0.045	-0.043	-0.045	-0.043
Sci-Fi & Fantasy	0.298	0.296	0.299	0.298	0.300	0.299
Netflix Original	0.957***	0.955***	0.959***	0.958***	0.955***	0.953***
logAv.Length	-0.272	-0.278	-0.273	-0.274	-0.270	-0.271
Constant	-11.581***	-12.726***	-11.598***	-11.971***	-11.808***	-11.993***
Observations	64,809	64,809	64,809	64,809	64,809	64,809
Pseudo R2	0.466	0.465	0.466	0.466	0.467	0.467

# Robustness checks: Excluding the US

Table: Robustness check: excluding the US

	Excl. as origin			Excl. as destination		Excl. as both
	(1) All	(2) All	(3) Foreign	(4) All	(5) Foreign	(6) Foreign
logDist.	-0.528***	-0.168	-0.043	-0.145**	-0.178***	-0.164**
Domestic		1.411***		0.767***		
logMigrants			0.538***		0.181***	0.399***
Common language	1.577***	1.515***	1.450***	1.573***	1.693***	0.448**
Contiguity	0.002	0.476*	-0.209	0.093	-0.148	0.016
Colony	-0.502	-0.864**	-0.865**	-0.140	-0.245**	0.246
logRating	12.986***	12.949***	12.813***	9.666***	9.589***	13.756***
logAge	-1.131***	-1.132***	-1.039***	-0.655***	-0.617**	-1.097***
Children & Family	-3.109***	-3.085***	-2.976***	-2.271***	-2.216***	-3.295***
Comedy	-0.980	-0.965	-1.058	0.201	0.222	-1.103
Crime & Thriller	-0.832	-0.825	-1.006*	-0.391	-0.414	-0.681
Drama	-1.755**	-1.735**	-1.895**	-0.672	-0.666	-1.672**
Horror	-0.492	-0.486	-0.531	-0.389**	-0.374**	-0.581
Romance	-2.522***	-2.505***	-4.004***	0.186	0.131	-3.498***
Sci-Fi & Fantasy	-0.202	-0.193	-0.276	0.366	0.377	-0.317
Netflix Original	2.367***	2.417***	2.721***	0.824***	0.810***	2.167***
logAv.Length	-0.879	-0.863	-0.851	-0.281	-0.265	-0.906
Constant	-9.645*	-12.542**	-19.625***	-7.882***	-9.541***	-18.469***
Observations	40,954	40,954	38,950	64,435	62,455	36,596
Pseudo R2	0.606	0.611	0.632	0.465	0.467	0.640

PPML estimation results. Dependant variable: number of clicks. All results include origin-time and destination-time fixed effects. Standard errors are clustered at the origin-time and destination-time level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

# Robustness checks: Title fixed effects

Table: Title fixed effects, TV Shows

	All		Foreign	
	(1)	(2)	(3)	(4)
logDist.	-0.522*** (0.053)	-0.186** (0.076)	-0.181** (0.081)	-0.257*** (0.064)
Domestic		1.124*** (0.237)		
logMigrants				0.419*** (0.093)
Common language	1.308*** (0.225)	1.484*** (0.174)	1.495*** (0.262)	1.464*** (0.220)
Contiguity	-0.674* (0.369)	-0.021 (0.286)	0.192 (0.222)	-0.492** (0.228)
Colony	-0.179 (0.168)	-0.215* (0.123)	-0.324*** (0.098)	-0.235** (0.096)
Constant	16.913*** (0.399)	13.717*** (0.687)	12.767*** (0.705)	8.689*** (1.040)
Observations	25,238	25,238	23,850	23,850
Pseudo R2	0.786	0.788	0.710	0.713

PPML estimation results. Dependant variable: number of clicks. All results include title, origin-time and destination-time fixed effects. Standard errors are clustered at the origin-time and destination-time level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

# Conclusion

- ▶ Fragmentation of Netflix catalogues.
- ▶ The title count of European content is below the proportion suggested by the European Commission.
- ▶ Gravity appears more important for clicks than for availability.
- ▶ Strong preference for home productions.
- ▶ Show/movie characteristics (quality) affect the number of clicks.
- ▶ The Netflix Original label is a major driver of clicks (more for movies than TV shows).
- ▶ For non-Netflix productions gravity matters even more.

Thank you for your attention!



Table: Summary statistics: TV Shows

Variable	TV Shows				
	Obs.	Mean	Std. Dev.	Min	Max
Clicks	68,432	10,892	237,853	0.00	33,400,000
logDist.	68,432	8.41	1.04	4.23	9.80
Domestic	68,432	0.05	0.21	0.00	1.00
Common language	68,432	0.17	0.37	0.00	1.00
Contiguity	68,432	0.06	0.24	0.00	1.00
Colony	68,432	0.11	0.31	0.00	1.00
Rating	68,432	7.25	1.01	1.70	9.50
logRating	68,432	1.97	0.16	0.53	2.25
Age	68,432	6.85	7.26	1.00	68.00
logAge	68,432	1.58	0.80	0.00	4.22
Average Length	68,432	36.88	17.91	1.00	120.00
logAv. Length	68,432	3.48	0.55	0.00	4.79
Netflix Original	68,432	0.45	0.50	0.00	1.00

Table: Summary statistics: Movies

Variable	Obs.	Movies			
		Mean	Std. Dev.	Min	Max
Clicks	135,256	30.95	1,904	0.00	394,453
logDist.	135,256	8.35	0.99	4.23	9.88
Domestic	135,256	0.06	0.24	0.00	1.00
Common language	135,256	0.22	0.41	0.00	1.00
Contiguity	135,256	0.08	0.27	0.00	1.00
Colony	135,256	0.14	0.34	0.00	1.00
Rating	135,256	6.17	1.08	1.40	9.30
logRating	135,256	1.80	0.20	0.34	2.23
Age	135,256	10.54	10.13	1.00	100.00
logAge	135,256	1.98	0.86	0.00	4.61
Length	135,256	101.68	20.73	2.00	251.00
logLength	135,256	4.60	0.25	0.69	5.53
Netflix Original	135,256	0.12	0.32	0.00	1.00

# Robustness checks: Interactions

Table: Netflix Original and Rating interaction

	(1)	(2)	(3)	(4)
logDist.	-0.334***	-0.579***	-0.296***	-0.169
Domestic	0.777***			1.408***
logMigrants			0.436***	
Common language	1.528***	1.423***	1.575***	1.515***
Contiguity	-0.189	-0.685	-0.596**	0.477*
Colony	-0.259	-0.238	-0.319***	-0.863**
logRating	24.033***	24.006***	19.279***	14.240***
logAge	-0.538**	-0.537**	-0.679***	-1.132***
Netflix Original	35.881***	35.818***	25.807***	5.318
logAv.Length	-0.335*	-0.334*	-0.228	-0.856
netrating	-16.618***	-16.592***	-11.821***	-1.385
Constant	-35.692***	-33.322***	-31.854***	-15.275**
Observations	68,432	68,432	64,809	40,954
Pseudo R2	0.549	0.548	0.481	0.611

PPML estimation results. Dependant variable: number of clicks. All results include origin-time, destination-time fixed effects. Standard errors are clustered at title level.  
 \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.